



SEQUENCE LISTING

<110> Tchaga, Grigory S.
Jokhadze, George

<120> Metal Ion Affinity Tags and Methods for
Using the Same

<130> CLON-056CIP

<140> US 09/858,332

<141> 2001-05-15

<150> 09/404,017

<151> 1999-09-23

<150> 60/101,867

<151> 1998-09-25

<160> 20

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 1

His Leu Ile His Asn Val His Lys Glu Glu His Ala His Ala His Asn
1 5 10 15

<210> 2

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 2

His Asp Asp His Asp Asp His Asp Asp His Asp Asp His
1 5 10 15
Asp Asp

<210> 3

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 3

His Glu Glu His Glu Glu His Glu Glu His Glu Glu His
1 5 10 15

<210> 4

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 4

His Asp Glu His Asp Glu His Asp Glu His Asp Glu His
1 5 10 15
Asp Glu

<210> 5

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> affinity peptide

<400> 5

His Glu Asp His Glu Asp His Glu Asp His Glu Asp His
1 5 10 15
Glu Asp

<210> 6

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> enterokinase cleavage site

<400> 6

Ile Glu Gly Arg
1

<210> 7

<211> 6

<212> PRT

<213> Artificial Sequence

<220>
<223> a factor Xa cleavage site

<400> 7
Leu Val Pro Arg Gly Ser
1 5

<210> 8
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> a thrombin cleavage site

<400> 8
His Pro Phe His Leu Val Ile His
1 5

<210> 9
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> a renin cleavage site

<400> 9
Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5 10

<210> 10
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> an immunological tag

<400> 10
Asp Tyr Lys Asp Asp Asp Lys
1 5

<210> 11
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> an immunological tag

<400> 11
Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1

5

10

<210> 12

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> an immunological tag

<400> 12

Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1

5

10

<210> 13

<211> 3430

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA sequence of vector containing cDNA of recombinant enterokinase

<400> 13

gacgaaaggg cctcgtgata cgcctatatt tataggttaa tgtcatgata ataatggttt 60
cttagacgtc aggtggcact ttccggggaa atgtgcgcgg aaccctatt tgtttatatt 120
tctaaatata ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat 180
aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgcctt attccctttt 240
ttgcggcatt ttgccttccg gtttttgctc acccagaaac gctggtgaaa gtaaaagatg 300
ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga 360
tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc 420
tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggc cgccgcatac 480
actattctca gaatgacttg gttgagtact caccagtac agaaaagcat cttacggatg 540
gcatgacagt aagagaatta tgcagtgtcg ccataaccat gaggatgata actgcggtcca 600
acttacttct gacaacgacg ggaggaccga aggagctaac cgcttttttg cacaacatgg 660
gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg 720
acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg 780
gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag 840
ttgcaggacc acttctgctc tcggcccttc cggctggctg gtttattgct gataaatctg 900
gagccgggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct 960
cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac 1020
agatcgtctga gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact 1080
catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga 1140
tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt 1200
cagaccccgat agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaactc 1260
gctgcttgca aacaaaaaaa ccaccgctac cagcgggtgg ttggttgccg gatcaagagc 1320
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc 1380
ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc 1440
tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg 1500
ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt 1560
cgtgcacaca gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg 1620
agctatgaga aagcgccacg cttcccgaag ggagaaagc ggacaggtat ccggtaaagc 1680
gbgcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct 1740
ttatagtctt gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc 1800
agggggggcgg agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt 1860

```

ttgctggcct tttgtcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg 1920
tattaccgcc tttgagttag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga 1980
gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg 2040
gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg gcagtgagcg 2100
caacgcaatt aatgtgagtt agctcactca ttaggcaccc caggctttac actttatgct 2160
tccggctcgt atgttgtgtg gaattgtgag cggataacaa tttcacacag gaaacagcta 2220
tgaccatgat tacgccaaagc ttgaaggatc atctcatcca caatgtccac aaagaggagc 2280
acgctcatgc ccacaacaag atcgatattg tcggaggaag tgactccaga gaaggagcct 2340
ggccttgggt cgttgtctctg tatttcgacg atcaacaggt ctgcggagct tctctggtga 2400
gcagggattg gctggtgtcg gccgccact gcgtgtacgg gagaaatag gagccgtcta 2460
agtggaaagc agtgctaggc ctgcatatgg catcaaactc gacttctcct cagatagaaa 2520
ctaggttgat tgaccaaatt gtcataaacc cacactacaa taaacggaga aagaacaatg 2580
acattgccat gatgcactct gaaatgaaag tgaactacac agattatata cagcctatatt 2640
gtttaccaga agaaaatcaa gtttttcccc caggaagaat ttgttctatt gctggctggg 2700
gggcacttat atatcaaggt tctactgcag acgtactgca agaagctgac gttccccttc 2760
tatcaaatga gaaatgtcaa caacagatgc cagaatataa cattacggaa aatatggtgt 2820
gtgcaggcta tgaagcagga ggggtagatt cttgtcaggg ggattcaggc ggaccactca 2880
tgtgccaaga aaacaacaga tggctcctgg ctggcgtgac gtcatttggg tatcaatgtg 2940
cactgcctaa tcgcccaggg gtgtatgccc ggggtcccaag gttcacagag tggatacaaa 3000
gttttctaca tgagctcgta attagctgag aattcactgg ccgtcgtttt acaacgtcgt 3060
gagcctggga aaacctggc gttacccaac ttaatcgctt tgcagcatat ccccttttcg 3120
ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc 3180
tgaatggcga atggcgctg atgcggtatt ttctccttac gcactgtgac ggtatttcac 3240
accgcatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagcccc 3300
gacaccgcc aacaccgct gacgcgccct gacgggcttg tctgctcccg gcattccgctt 3360
acagacaagc tgtgaccgtc tccgggagct gcattgtgtca gaggttttca ccgtcatcac 3420
cgaaacgcgc                                     3430

```

<210> 14

<211> 279

<212> PRT

<213> Artificial Sequence

<220>

<223> protein sequence of vector containing cDNA of recombinant enterokinase

<400> 14

Met Thr Met Ile Thr Pro Ser Leu Lys Asp His Leu Ile His Asn Val
1 5 10 15
His Lys Glu Glu His Ala His Ala His Asn Lys Ile Asp Ile Val Gly
20 25 30
Gly Ser Asp Ser Arg Glu Gly Ala Trp Pro Trp Val Val Ala Leu Tyr
35 40 45
Phe Asp Asp Gln Gln Val Cys Gly Ala Ser Leu Val Ser Arg Asp Trp
50 55 60
Leu Val Ser Ala Ala His Cys Val Tyr Gly Arg Asn Met Glu Pro Ser
65 70 75 80
Lys Trp Lys Ala Val Leu Gly Leu His Met Ala Ser Asn Leu Thr Ser
85 90 95
Pro Gln Ile Glu Thr Arg Leu Ile Asp Gln Ile Val Ile Asn Pro His
100 105 110
Tyr Asn Lys Arg Arg Lys Asn Asn Asp Ile Ala Met Met His Leu Glu
115 120 125
Met Lys Val Asn Tyr Thr Asp Tyr Ile Gln Pro Ile Cys Leu Pro Glu
130 135 140
Glu Asn Gln Val Phe Pro Pro Gly Arg Ile Cys Ser Ile Ala Gly Trp

145 150 155 160
 Gly Ala Leu Ile Tyr Gln Gly Ser Thr Ala Asp Val Leu Gln Glu Ala
 165 170 175
 Asp Val Pro Leu Leu Ser Asn Glu Lys Cys Gln Gln Gln Met Pro Glu
 180 185 190
 Tyr Asn Ile Thr Glu Asn Met Val Cys Ala Gly Tyr Glu Ala Gly Gly
 195 200 205
 Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Glu
 210 215 220
 Asn Asn Arg Trp Leu Leu Ala Gly Val Thr Ser Phe Gly Tyr Gln Cys
 225 230 235 240
 Ala Leu Pro Asn Arg Pro Gly Val Tyr Ala Arg Val Pro Arg Phe Thr
 245 250 255
 Glu Trp Ile Gln Ser Phe Leu His Glu Leu Val Ile Ser Glu Phe Thr
 260 265 270
 Gly Arg Arg Phe Thr Thr Ser
 275

<210> 15
 <211> 12
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> an amino acid sequence embodiment of the affinity
 purification site

<400> 15
 His Asn His Asn His Asn His Asn His Asn His Asn
 1 5 10

<210> 16
 <211> 48
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> a DNA sequence embodiment of the affinity
 purification site

<400> 16
 catctcatcc acaatgtcca caaagaggag cacgctcatg cccacaac 48

<210> 17
 <211> 36
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> a DNA sequence embodiment of the affinity
 purification site

<400> 17
 cataaccata accataacca taaccataac cataac 36

<210> 18
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> a DNA sequence embodiment of the affinity
purification site

<400> 18
catgatgac atgatgatca tgatgatcat gatgatcatg atgatcatga tgat 54

<210> 19
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> a DNA sequence embodiment of the affinity
purification site

<400> 19
catgaggagc atgaggagca tgaggagcat gaggagcatg aggagcatga ggag 54

<210> 20
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> a DNA sequence embodiment of the affinity
purification site

<400> 20
catgatgagc atgatgagca tgagaacat gagaacatg aggatcatga ggat 54

a
ant